Tests of bivariate stochastic order

Yunfeng Liu

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of Master of Science in Mathematics [1]

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

[1] The M.Sc. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

Abstract

The purpose of this thesis is to compare rank-based tests of bivariate stochastic order. Given two bivariate distributions F and G, the general problem we are dealing with is to test $H_0 : F = G$ against $H_1 : F < G$, where F and G are independent continuous distributions on $\Re^2$. ("$F < G$" means that $F(x) \leq G(x)\ \forall x \in \Re^2$, and $\exists x \in \Re^2$ such that $F(x) < G(x)$.). In particular, we will analyze three analogues of the one-dimensional Mann-Whitney-Wilcoxon test in two dimensions. Two of the test statistics are new; we call them the Kendall and Spearman statistics. We will then show the asymptotic distributions and carry out empirical comparisons of the Kendall, Spearman and the third two-dimensional Mann-Whitney-Wilcoxon statistics.

Acknowledgements

First and foremost I am heartily thankful to my supervisor Professor Gail Ivanoff, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. Her way of removing misunderstanding gave me the opportunity to pursue this work in the best conditions. One simply could not wish for a better or friendlier supervisor. Thank you deeply.

I also offer my regards to all of those who supported me in any respect throughout the process of writing this thesis, such as Mr. Jiang Mei, Mr. Jesse Collingwood, and so on.

Lastly, I wish to express my love and gratitude to my beloved wife Siwei and little son Oscar for their understanding and endless love throughout the duration of my studies. Thank you for my parents for continuous support, selfishless love and encouragement. Thank you for my friend Yi Zhan for his understanding and supporting.

Dedication

This work is dedicated to my son Oscar Liu, to my lovely wife Siwei Luo.

Contents

List of Tables

Chapter 1

Introduction

The purpose of this thesis is to compare rank-based tests of bivariate stochastic order. Given two bivariate distributions F and G, the general problem we are dealing with is to test $H_0 : F = G$ against the alternative $H_1 : F < G$, where F and G are independent continuous distributions on $\Re^2$. ("$F < G$" means that $F(x) \leq G(x)$ $\forall x \in \Re^2$, and $\exists x \in \Re^2$ such that $F(x) < G(x)$.). In particular, we will analyze analogues of the one-dimensional Mann-Whitney-Wilcoxon test in two dimensions.

For distributions F and G on $\Re$, there are various distribution free tests of stochastic order, including the Kolmogorov-Smirnov test, the Mann-Whitney-Wilcoxon test, and tests based on P-P plots. They will be discussed briefly in Section 1.1.

For distributions on $\Re^2$, the problem becomes much more complicated as none of the natural analogues of the one-dimensional test statistics are distribution free. Since the Mann-Whitney-Wilcoxon statistic performs best for distributions on $\Re$, we will focus on generalizing this statistic. We first consider a well known statistic that is based on comparisons of marginal distributions of F and G. However, the resulting test is based on the assumption that G is a shifted version of F. To remove this restriction, we will introduce two new statistics which also generalize the classical one-dimensional Mann-Whitney-Wilcoxon test in bivariate form. We will call them the Kendall statistic and Spearman statistic respectively. Also, we will state and prove some theorems about the asymptotic behaviour of these statistics. Finally, we will compare their performance empirically.

1.1 The test statistics in $\Re$

Assume that F and G are distributions on $\Re$ and that each have positive densities f and g on their open support. Let $X_1, ..., X_n$ be i.i.d. F and $Y_1, ..., Y_m$ be i.i.d G. Let F_n and G_m denote the respective empirical distributions:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le x\}} \tag{1.1.1}$$

$$G_m(x) = \frac{1}{m}\sum_{j=1}^{m} I_{\{Y_j \le x\}} \tag{1.1.2}$$

We start by reviewing three well known non-parametric tests of $H_0 : F = G$ vs. $H_1 : F < G$

1. **Kolmogorov-Smirnov test statistic**

   The Kolmogorov-Smirnov test statistic is defined as:

$$D_{n,m} = \sup_x \left[G_m(x) - F_n(x)\right]. \tag{1.1.3}$$

   If $F = G$, then as $n, m \to \infty$ ([3], Section 3.7)

$$\sqrt{\frac{nm}{n+m}} D_{n,m} \Rightarrow \sup_{0<p<1} B(p)$$

where B is a Brownian bridge on $[0,1]$. and '$\Rightarrow$' denotes convergence in distribution.

A Brownian bridge on $[0,1]$ is defined by:

$$B(p) = W(p) - W(1) \quad p \in [0,1] \tag{1.1.4}$$

where $W(p)$ is a standard Wiener process. In particular, B is a continuous Gaussian process on $[0,1]$ with covariance function $Cov(B(p),B(q)) = p \wedge q - pq$ for $0 \leq p,q \leq 1$, where $p \wedge q = \min(p,q)$.

Furthermore,

$$P\left(\sup_{0<p<1} B(p) > x\right) = e^{-2x^2},$$

so for a test of level approximately α, reject H_0 if $\sqrt{\frac{nm}{n+m}} D_{n,m} > \sqrt{-\frac{\ln \alpha}{2}}$.

2. **Mann-Whitney-Wilcoxon test statistic (Abrev. M-W-W)**

The Mann-Whitney version of this test statistic is:

$$U_{n,m} = \int_{-\infty}^{\infty} G_m(x) dF_n(x) = \frac{1}{n}\sum_{i=1}^{n} G_m(X_i) = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} I_{\{Y_j \leq X_i\}}. \tag{1.1.5}$$

If $F = G$, then as $n,m \to \infty$

$$\frac{(U_{n,m} - 1/2)}{\sqrt{\frac{(n+m+1)}{12nm}}} \Rightarrow Z,$$

where $Z \sim N(0,1)$. Therefore, for a test of level approximately α, reject H_0 if

$$\frac{(U_{n,m} - 1/2)}{\sqrt{\frac{(n+m+1)}{12nm}}} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} I_{\{Y_j \leq X_i\}} - nm/2}{\sqrt{nm(n+m+1)/12}} > z_\alpha.$$

3. **P-P plot statistic**

Given two distributions F and G on $\Re$, the $P-P$ plot of G against F is defined to be $G(F^{-}(p))$ for $0 \leq p \leq 1$ ($F^{-}(p) = \inf(x : F(x) \geq p)$ for $0 < p \leq 1$ and

$F^-(0) = \lim_{p\to 0} F^-(p)$). $G_m(F_n^-(p))$ is the empirical P-P plot of G_m against F_n.

The P-P plot statistic is defined as:

$$W_{n,m} = \sup_{0<p<1} \left[G_m(F_n^-(p)) - p\right] \tag{1.1.6}$$

If $k = \lceil np \rceil$, the smallest integer greater than or equal to np, then $F_n^-(p) = X_{(k)}$, the k^{th} order statistic corresponding to $X_1, ..., X_n$. Therefore, if $k = \lceil np \rceil$,

$$G_m(F_n^-(p)) = \frac{1}{m}\sum_{j=1}^{m} I_{\{Y_j \le X_{(k)}\}}.$$

If $F = G$, then as $n, m \to \infty$ ([3], Section 3.9)

$$\sqrt{\frac{nm}{n+m}} W_{n,m} \Rightarrow \sup_{0<p<1} B(p)$$

where B is a Brownian bridge on $[0, 1]$.
As before,

$$P\left(\sup_{0<p<1} B(p) > x\right) = e^{-2x^2},$$

so for a test of level approximately α, we reject H_0 if $\sqrt{\frac{nm}{n+m}} W_{n,m} > \sqrt{-\frac{\ln \alpha}{2}}$.

In fact, it is easily seen that $D_{n,m} \approx W_{n,m}$, so the P-P plot statistic is asymptotically equivalent to the Kolmogorov-Smirnov test statistic. In section 4.1 we compare these three statistics empirically and conclude that Mann-Whitney-Wilcoxon has more power than the others.

1.2 The bivariate case

The three tests in section 1.1 are all distribution free under H_0, which is because $F(X)$ and $G(Y)$ are uniformly distributed. However, this will not be true anymore in the bivariate case. We will focus on analogues of the Mann-Whitney-Wilcoxon test in the bivariate case and its derived offsprings; the Kendall and Spearman statistics. All of these results can be generalized to d-dimensions, but for clarity we will focus on $d = 2$.

Chapter 2

Bivariate tests

Let F and G be distribution functions on $\Re^2$ such that each have positive densities f and g on their open support. Let F_1, F_2 and G_1, G_2 be the marginal distributions of F and G. If $X_1, X_2, ...X_n$ and $Y_1, Y_2, ...Y_m$ are iid from F and G respectively, and $N = n + m$, then the empirical distribution functions F_n, G_m, H_N are defined as:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_{i1}\le x_1, X_{i2}\le x_2\}} \tag{2.0.1}$$

$$G_m(x) = \frac{1}{m}\sum_{i=1}^{m} I_{\{Y_{i1}\le x_1, Y_{i2}\le x_2\}} \tag{2.0.2}$$

$$H_N(x) = \frac{1}{N}\sum_{i=1}^{N} I_{\{Z_{i1}\le x_1, Z_{i2}\le x_2\}} \tag{2.0.3}$$

where $x = (x_1, x_2)$, $X_i = (X_{i1}, X_{i2})$, $Z = (Z_1, Z_2, \ldots, Z_N)$ is the combined sample of $X_1, ..X_n$ and $Y_1, ...Y_m$. The empirical marginal distributions $F_n^1, F_n^2, G_m^1, G_m^2, H_N^1, H_N^2$ are all defined in the same way as the one-dimensional empirical distribution function. For instance, $H_N^1(x) = \frac{1}{N}\sum_{i=1}^{N} I_{\{Z_{i1}\le x\}}$.

2.1 Marginal Mann-Whitney-Wilcoxon test

The multivariate Mann-Whitney-Wilcoxon test is well described in [1], and we briefly present it here. Since this statistic is based on each of the components of the random vectors X_i, we call it the marginal Mann-Whitney-Wilcoxon test.

The marginal Mann-Whitney-Wilcoxon test statistic assumes that G is a shifted version of F, which is not always reasonable in practice. To be precise, it is assumed that $G(x) = F(x+\Delta)$, where $\Delta = (\Delta_1, \Delta_2)$, Δ_1, Δ_2 could be any value in $\Re$. We will also see in section 4 that if F and G are two completely different distributions but with the same marginal distributions, then the marginal Mann-Whitney-Wilcoxon test statistic won't reject the null hypothesis even though it should. Moreover, another drawback of the marginal Mann-Whitney-Wilcoxon test statistic is that it tests the alternative $H_1 : \Delta \neq 0$.

Define R_{ti} to be the rank of $X_t = (X_{t,1}, X_{t,2}), t = 1 \dots n$, in the combined sample of the ith component ($i \in 1, 2$). Let $N = n + m$, and define U_i as:

$$U_i = \sum_{t=1}^{n} \left[\frac{R_{ti}}{N+1} - \frac{1}{2} \right]. \tag{2.1.1}$$

Let $\mathbf{U} = (U_1, U_2)$, and assume that as $n, m \to \infty$, $\frac{m}{n+m} \to \lambda$, $0 < \lambda < 1$. Then under the $H_0 : \Delta = 0$,

$$\frac{1}{\sqrt{n+m}} \mathbf{U} \Rightarrow N(0, \mathbf{V}).$$

$N(0, \mathbf{V})$ is bivariate normal with covariance matrix $\mathbf{V} = ((v_{ij})), i, j = 1, 2$, where

$$v_{ij} = \frac{\lambda(1-\lambda)}{12} \quad i = j \tag{2.1.2}$$

$$v_{ij} = \lambda(1-\lambda) \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i(x) F_j(y) dF_{ij}(x, y) - \frac{1}{4} \right\} \quad i \neq j. \tag{2.1.3}$$

In [1], the resulting test statistic is :

$$U^* = \frac{1}{N}\mathbf{U}'\left(\hat{\mathbf{V}}\right)^{-1}\mathbf{U} \tag{2.1.4}$$

where $\hat{V} = ((\hat{v}_{ij}))$ estimates $\mathbf{V}$, U' is the transpose of U. The components $\hat{v}_{ij}$ have the following form:

$$\hat{v}_{ii} = Var\left(\frac{1}{\sqrt{N}}U_i\right) = \frac{mn}{12N(N+1)} \tag{2.1.5}$$

$$\hat{v}_{ij} = \frac{mn}{N^2(N-1)(N+1)^2}\left(\sum_{t=1}^{N} R_{ti}R_{tj} - \frac{N(N+1)^2}{4}\right) \tag{2.1.6}$$

It is been proven in [1] that under H_0, U^* is asymptotically chi-square with 2 degrees of freedom.

2.2 Kendall and Spearman Statistics

In order to generalize the shift assumption, here we propose two alternatives to the marginal Mann-Whitney-Wilcoxon statistic. We will consider their asymptotic properties in the next section. For F_n, G_m and H_N as in (2.0.1)-(2.0.3), define the empirical integrals of G_m with respect to H_N as follows (The empirical integral of F_n with respect to H_N can be defined similarly.):

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_m dH_N \\ = \;& \frac{1}{N}\frac{1}{m}\sum_{i=1}^{N}\sum_{j=1}^{m} I(Y_{j1} \leq Z_{i1}, Y_{j2} \leq Z_{i2}). \end{aligned} \tag{2.2.1}$$

Define another empirical integral of G_m with respect to $H_N^1 \times H_N^2$ as follows:

$$\begin{aligned} &\int_{\Re^2} G_m dH_N^1 dH_N^2 \\ = \;& \frac{1}{N^2}\frac{1}{m}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{h=1}^{m} I(Y_{h1} \leq Z_{i1})I(Y_{h2} \leq Z_{j2}). \end{aligned} \tag{2.2.2}$$

2.2.1 Kendall statistic

The Kendall statistic is defined as follows:

$$K_{nm} = \sqrt{\frac{nm}{n+m}} \int_{\Re^2} (G_m - F_n) dH_N. \tag{2.2.3}$$

Rewrite K_{nm} in terms of summations:

$$\begin{aligned} K_{nm} &= \sqrt{\frac{nm}{n+m}} \int_{\Re^2} (G_m - F_n) dH_N \\ &= \sqrt{\frac{nm}{n+m}} \left(\int_{\Re^2} G_m dH_N - \int_{\Re^2} F_n dH_N \right) \\ &= \sqrt{\frac{nm}{n+m}} \left[\frac{1}{N} \frac{1}{m} \sum_{i=1}^{N} \sum_{j=1}^{m} I(Y_{j1} \le Z_{i1}, Y_{j2} \le Z_{i2}) \right] \\ & -\sqrt{\frac{nm}{n+m}} \left[\frac{1}{N} \frac{1}{n} \sum_{i=1}^{N} \sum_{j=1}^{n} I(X_{j1} \le Z_{i1}, X_{j2} \le Z_{i2}) \right]. \end{aligned} \tag{2.2.4}$$

We call it the Kendall statistic because Kendall's tau is defined as:

$$\tau^F = 4 \int \int F(x,y) dF(x,y) - 1 \tag{2.2.5}$$

and when $H_0 : F = G$ is true, F_n, G_m and H_N are all estimators of F. Therefore K_{nm} compares estimators of τ^F.

2.2.2 Spearman statistic

The Spearman statistic is defined as follows:

$$S_{nm} = \sqrt{\frac{nm}{n+m}} \int_{\Re^2} (G_m - F_n) dH_N^1 dH_N^2. \tag{2.2.6}$$

Rewrite S_{nm} in terms of summations:

$$
\begin{aligned}
S_{nm} &= \sqrt{\frac{nm}{n+m}} \int_{\Re^2} (G_m - F_n) dH_N^1 dH_N^2 && (2.2.7) \\
&= \sqrt{\frac{nm}{n+m}} \left(\int_{\Re^2} G_m dH_N^1 dH_N^2 - \int_{\Re^2} F_n dH_N^1 dH_N^2 \right) \\
&= \sqrt{\frac{nm}{n+m}} \left[\frac{1}{m} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{h=1}^{m} I(Y_{h1} \le Z_{i1}) I(Y_{h2} \le Z_{j2}) \right] \\
&\quad - \sqrt{\frac{nm}{n+m}} \left[\frac{1}{n} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{h=1}^{n} I(X_{h1} \le Z_{i1}) I(X_{h2} \le Z_{j2}) \right]. && (2.2.8)
\end{aligned}
$$

We call it the Spearman statistic because Spearman's rho is defined as follows:

$$\rho^F = 12 \int \int F(x, y) dF_1(x) dF_2(y) - 3 \qquad (2.2.9)$$

and when $H_0 : F = G$ is true, S_{nm} compares estimators of ρ^F.

We will consider the asymptotic behaviors of the Kendall and Spearman statistics in the next chapter.

Chapter 3

Asymptotic Distributions

If $H_1 : F < G$ is true, K_{nm} and S_{nm} should get large. In this chapter, we find the limiting distributions of K_{nm} and S_{nm} when H_0 is true so that appropriate critical values can be found.

3.1 Limiting Distributions

We begin by reviewing some facts about copulas (for details see [2]).

Definition 3.1.1 (Definition 2.2.2 in [2]) *A two-dimensional copula is a function* $C : [0,1]^2 \to [0,1]$ *with the following properties:*

1. *Domain of C is* $[0,1]^2$

2. $\forall u, v \in [0,1]$

$$C(u,0) = 0 = C(0,v)$$

$$C(u,1) = u, C(1,v) = v$$

3. *For every* u_1, u_2, v_1, v_2 *in* $[0,1]$ *such that* $u_1 \le u_2$, $v_1 \le v_2$,

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \ge 0$$

With an appropriate extension of its domain to $\Re^2$, every copula is a joint distribution function with marginals that are uniform on $[0,1]$.

Theorem 3.1.2 ***Sklar's theorem***:*[Theorem 2.3.3 in [2]] Let F be a joint distribution function with marginals F_1 and F_2. Then there exists a copula C such that for all x_1, x_2 in $\bar{\Re}$*

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \tag{3.1.1}$$

If F_1 and F_2 are continuous, then C is unique; otherwise, C is uniquely determined on $Range(F_1) \times Range(F_2)$. Conversely, if C is a copula and F_1, F_2 are distribution functions, then the function F defined above is a joint distribution function with marginals F_1 and F_2.

Hence by Sklar's theorem, for F a continuous distribution on $\Re^2$ with marginals F_1, F_2, $C^F(p) := F(F_1^-(p_1), F_2^-(p_2)), p \in [0,1]^2$ is a copula, and for all $x \in \Re^2$, $F(x) = C^F(F_1(x_1), F_2(x_2))$. Also, given a sample $X_1, X_2, \ldots, X_n$ from a distribution F, the empirical copula is defined by

$$C_n^F(p_1, p_2) = F_n(F_n^{1-}(p_1), F_n^{2-}(p_2))$$

.

Definition 3.1.3 (cf. [3]) *For H a distribution on $\Re^2$, we denote the associated Brownian bridge on $\Re^2$ by B^H: i.e. B^H is a mean zero Gaussian process with $Cov(B^H(s), B^H(t)) = H(s \wedge t) - H(s)H(t)\ \forall\ s, t \in \Re^d$. (For $s = (s_1, s_2)$, $t = (t_1, t_2)$, $s \wedge t = (s_1 \wedge t_1, s_2 \wedge t_2)$.)*

Definition 3.1.4 (Definition 2.1 in [4]) *Let F_1^-, F_2^- be the inverses of marginal distributions F_1, F_2. The two-dimensional P-P plot of G against F is defined as $G(F_1^-, F_2^-)$, where for $p = (p_1, p_2) \in [0,1]^2$,*

$$G(F_1^-, F_2^-)(p_1, p_2) := G(F_1^-(p_1), F_2^-(p_2)) \tag{3.1.2}$$

Also, the empirical $P-P$ plot of G against F is the $P-P$ plot of G_m against F_n: $G_m(F_n^{1-}, F_n^{2-})$.

Now consider the empirical P-P plots of F_n and G_m against H_N. Let $\psi_{n,m}$ be defined as:

$$\begin{aligned} &\psi_{n,m}(p_1, p_2) \\ &:= \sqrt{\frac{nm}{n+m}} \left[G_m(H_N^{1-}(p_1), H_N^{2-}(p_2)) - F_n(H_N^{1-}(p_1), H_N^{2-}(p_2)) \right]. \end{aligned} \tag{3.1.3}$$

Theorem 3.1.5 *([4], Theorem 3.8). Under $H_0: F = G$, as $n, m \to \infty$ and $\frac{m}{n+m} \to \lambda$*

$$\psi_{n,m}(p_1, p_2) \Rightarrow B^{C^F} \tag{3.1.4}$$

in $\ell^\infty([0,1]^2)$(the space of almost surely bounded functions $f : [0,1]^2 \to \Re$ equipped with the sup norm).

Here $B^{C^F} \in \ell^\infty([0,1]^2)$ is because B^{C^F} is a.s. continuous and equal to 0 at $(0,0)$ and $(1,1)$.

To apply Theorem 3.1.5, note that we can rewrite the Kendall and Spearman statistics as follows:

$$\begin{aligned} K_{nm} &:= \sqrt{\frac{nm}{n+m}} \int_{\Re^2} (G_m - F_n) dH_N \\ &= \sqrt{\frac{nm}{n+m}} \int_{[0,1]^2} (G_m(H_N^{1-}(p_1), H_N^{2-}(p_2)) \\ &\quad - F_n(H_N^{1-}(p_1), H_N^{2-}(p_2))) dH_N(H_N^{1-}(p_1), H_N^{2-}(p_2)) \\ &= \int_{[0,1]^2} \psi_{n,m}(p_1, p_2) dC_N^H(p_1, p_2). \end{aligned} \tag{3.1.5}$$

$$
\begin{aligned}
S_{nm} &:= \sqrt{\frac{nm}{n+m}} \int_{\Re^2} (G_m - F_n) dH_N^1 dH_N^2 \\
&= \sqrt{\frac{nm}{n+m}} \int_{[0,1]^2} (G_m(H_N^{1-}(p_1), H_N^{2-}(p_2)) \\
&\quad - F_n(H_N^{1-}(p_1), H_N^{2-}(p_2))) dp_1, dp_2 \\
&= \int_{[0,1]^2} \psi_{n,m}(p_1, p_2) dp_1 dp_2. \qquad (3.1.6)
\end{aligned}
$$

Theorem 3.1.6 *Assume that $H_0 : F = G$ is true. As $n, m \to \infty$, $\frac{m}{n+m} \to \lambda$,*

$$S_{nm} \Rightarrow N(0, \sigma_s^2) \qquad (3.1.7)$$

where σ_s^2 is equal to:

$$
\begin{aligned}
\sigma_s^2 &= Var\left[\int_{[0,1]^2} B^{C^F} dp_1 dp_2\right] \\
&= \int_{[0,1]^2} \int_{[0,1]^2} C^F(p \wedge q) dp_1 dp_2 dq_1 dq_2 - \left(\int_{[0,1]^2} C^F(p_1, p_2) dp_1 dp_2\right)^2. \qquad (3.1.8)
\end{aligned}
$$

Proof: For $F \in \ell^\infty([0,1]^2)$, $T : F \mapsto \int_{[0,1]^2} F dp_1 dp_2$ is a bounded linear functional, because $\|T\| = \sup_{\|F\| \le 1} |T(F)| \le 1$. Since $\psi_{n,m} \in \ell^\infty([0,1]^2)$, and $S_{nm} = T(\psi_{n,m})$, by the continuous mapping theorem and Theorem (3.1.5), $S_{nm} \Rightarrow \int_{[0,1]^2} B^{C^F} dp_1 dp_2$. Let $Z = \int_{[0,1]^2} B^{C^F} dp_1 dp_2$. Since the mapping T is linear and B^{C^F} is Gaussian, Z is a Gaussian random variable. Letting $p = (p_1, p_2)$, $q = (q_1, q_2)$,

$$
\begin{aligned}
E[Z] &= \int_{[0,1]^2} E(B^{C^F}) dp_1 dp_2 \\
&= \int_{[0,1]^2} 0 dp_1 dp_2 = 0. \qquad (3.1.9)
\end{aligned}
$$

$$\begin{aligned}
\sigma_s^2 &= Var[Z] \\
&= E\left[\left(\int_{[0,1]^2} B^{C^F} dp_1 dp_2\right)^2\right] \\
&= E\left[\int_{[0,1]^2}\int_{[0,1]^2} B^{C^F}(p_1,p_2)B^{C^F}(q_1,q_2)dp_1dp_2dq_1dq_2\right] \\
&= \int_{[0,1]^2}\int_{[0,1]^2} E\left[B^{C^F}(p_1,p_2)B^{C^F}(q_1,q_2)\right] dp_1dp_2dq_1dq_2 \qquad (3.1.10) \\
&= \int_{[0,1]^2}\int_{[0,1]^2} Cov\left[B^{C^F}(p_1,p_2)B^{C^F}(q_1,q_2)\right] dp_1dp_2dq_1dq_2 \\
&= \int_{[0,1]^2}\int_{[0,1]^2} \left[C^F(p\wedge q) - C^F(p)C^F(q)\right] dp_1dp_2dq_1dq_2 \\
&= \int_{[0,1]^2}\int_{[0,1]^2} C^F(p\wedge q)dp_1dp_2dq_1dq_2 - \left(\int_{[0,1]^2} C^F(p)dp_1dp_2\right)^2. \qquad (3.1.11)
\end{aligned}$$

In equation (3.1.10), Fubini's Theorem applies, because the integrand $B^{C^F}(p_1,p_2)B^{C^F}(q_1,q_2)$ is almost surely continuous and bounded.

∎

To determine the limiting distribution of K_{nm}, we need the following Glivenko-Cantelli theorem:

Theorem 3.1.7 *[[4], Theorem 3.1] When $H_0 : F = G$ is true, as $n, m \to \infty$,*

$$\sup_{p\in[0,1]^2} \left|C_N^H(p) - C^F(p)\right| \to 0 \quad a.s.. \qquad (3.1.12)$$

Theorem 3.1.8 *Assume that $H_0 : F = G$ is true. As $n, m \to \infty$, $\frac{m}{n+m} \to \lambda$,*

$$K_{nm} \Rightarrow N(0, \sigma_k^2) \qquad (3.1.13)$$

where σ_k^2 is equal to:

$$\begin{aligned}
\sigma_k^2 &= Var\left[\int_{[0,1]^2} B^{C^F} dC^F(p)\right] \\
&= \int_{[0,1]^2}\int_{[0,1]^2} C^F(p\wedge q)dC^F(p)dC^F(q) - \left(\int_{[0,1]^2} C^F(p)dC^F(p)\right)^2 \qquad (3.1.14)
\end{aligned}$$

Proof: Under $H_0 : F = G$, $\psi_{n,m} \Rightarrow B^{C^F}$ by Theorem 3.1.5, and $\sup_{x\in[0,1]^2} |C_N^H - C^F| \to 0$ a.s. by the Glivenko-Cantelli theorem for the empirical copula, Theorem 3.1.7. Since B^{C^F} is continuous almost surely, we can apply Skorohod's representation theorem (see Theorem 6.7 in [5]). Using the same notation for our processes, there exists another probability space on which $(\psi_{n,m}, C_N^H) \to (B^{C^F}, C^F)$ a.s.. We may assume without loss of generality that on this space, B^{C^F} is continuous for all ω. We will use this representation to show that the asymptotic behavior of K_{nm} is equivalent to that of $\int \psi_{n,m} dC^F$.

Fix ω. Since B^{C^F} is continuous on $[0,1]^2$, $M = \sup_{[0,1]^2} |B^{C^F}| < \infty$, and $\exists$ a partition $0 = s_0 < s_1 < \ldots < s_h = 1$ and $0 = t_0 < t_1 < \ldots < t_k = 1$ such that $\sup_{p,q\in R_{i,j}} |B^{C^F}(p) - B^{C^F}(q)| < \epsilon$ where $R_{i,j} = [s_{i-1}, s_i) \times [t_{j-1}, t_j)$, $i = 1 \ldots h-1$, $j = 1 \ldots k-1$, and the intervals are closed on the right when $i = h$ or $j = k$. Define the discrete approximation $B = B^{C^F}(s_{i-1}, t_{j-1})$ on $R_{i,j}$, $i = 1, \ldots, h$ and $j = 1, \ldots, k$, and note that $\left\| B^{C^F} - B \right\| \leq \epsilon$. ($\|\cdot\|$ denotes the sup norm). Choose n, m sufficiently large such that $\left\| \psi_{n,m} - B^{C^F} \right\| \leq \epsilon$ and $\left\| C_N^F - C^F \right\| \leq \frac{\epsilon}{4hkM}$. (Recall that ω is fixed and so M, n and m may depend on ω.).

Now consider $\int \psi_{n,m} dC_N^H - \int \psi_{n,m} dC^F$:

$$
\begin{aligned}
\left| \int \psi_{n,m} d(C_N^H - C^F) \right| &\leq \left| \int (\psi_{n,m} - B^{C^F}) d(C_N^H - C^F) \right| \quad (a) \\
&+ \int \left\| B^{C^F} - B \right\| d(C_N^H - C^F) \quad (b) \\
&+ \left\| \int B d(C_N^H - C^F) \right\| \quad (c)
\end{aligned}
$$

$$
\begin{aligned}
(a) &\leq \int \epsilon dC_N^H + \int \epsilon dC^F = 2\epsilon. \\
(b) &\leq \int \epsilon C_N^H + \int \epsilon dC^F = 2\epsilon.
\end{aligned}
$$

For (c), since $(C_N^H - C^F)(R_{i,j}) = (C_N^H - C^F)(s_i, t_j) - (C_N^H - C^F)(s_{i-1}, t_j) - (C_N^H - C^F)(s_i, t_{j-1}) + (C_N^H - C^F)(s_{i-1}, t_{j-1})$, hence $(C_N^H - C^F)(R_{i,j}) \leq 4 \sup_{[0,1]^2} |C_N^H - C^F|$, Therefore,

$$\begin{aligned}(c) &= \left\| \sum_{i=1}^{h} \sum_{j=1}^{k} B^{C^F}(s_{i-1}, t_{j-1})((C_N^H - C^F)(R_{i,j})) \right\| \\ &\leq M4hk \ \left\| C_N^H - C^F \right\| \leq \epsilon. \end{aligned} \tag{3.1.15}$$

Hence, on this space, $\int \psi_{n,m} d(C_N^H - C^F) \to 0$ for almost all ω, as $n, m \to \infty$. Therefore, in general, under $H_0 : F = G$, by linearity of the map $B^{C^F} \to \int B^{C^F} dC^F$

$$\begin{aligned}\int \psi_{n,m} dC_N^H = \int \psi_{n,m} dC^F + \int \psi_{n,m} d(C_N^H - C^F) \\ \Rightarrow \int B^{C^F} dC^F =_D N(0, \sigma_k^2). \end{aligned} \tag{3.1.16}$$

Let $Z = \int B^{C^F} dC^F$. Since $E(Z) = 0$,

$$\begin{aligned}\sigma_k^2 &= Var\left[Z\right] \\ &= E\left[\left(\int_{[0,1]^2} B^{C^F} dC^F\right)^2\right] \\ &= E\left[\int_{[0,1]^2}\int_{[0,1]^2} B^{C^F}(p) B^{C^F}(q) dC^F(p) dC^F(q)\right] \\ &= \int_{[0,1]^2}\int_{[0,1]^2} E\left[B^{C^F}(p) B^{C^F}(q)\right] dC^F(p) dC^F(q) \\ &= \int_{[0,1]^2}\int_{[0,1]^2} Cov\left[B^{C^F}(p) B^{C^F}(q)\right] dC^F(p) dC^F(q) \\ &= \int_{[0,1]^2}\int_{[0,1]^2} (C^F(p \wedge q) - C^F(p) C^F(q)) dC^F(p) dC^F(q) \\ &= \int_{[0,1]^2}\int_{[0,1]^2} C^F(p \wedge q) dC^F(p) dC^F(q) - \left(\int_{[0,1]^2} C^F(p) dC^F(p)\right)^2 \end{aligned} \tag{3.1.17}$$

∎

3.2 Estimation of variances

By Sklar's theorem, the variances of σ_s^2 and σ_k^2 can be rewritten by using F, F_1 and F_2:

$$\begin{aligned}\sigma_k^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u\wedge t)dF(u)dF(t) - \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u)dF(u)\right]^2 \\ \sigma_s^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u\wedge t)dF_1(u)dF_2(u)dF_1(t)dF_2(t) \\ &\quad - \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u)dF_1(u)dF_2(u)\right]^2\end{aligned}$$

Under H_0, H_N converges uniformly to F almost surely, and so σ_k^2 and σ_s^2 can be estimated using their empirical counterparts:

$$\begin{aligned}\hat{\sigma}_k^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_N(u\wedge t)dH_N(u)dH_N(t) \\ &\quad - \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_N(u)dH_N(u)\right]^2 \qquad (3.2.1) \\ &= \frac{1}{N^3}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} I(Z_{k1}\le Z_{i1}\wedge Z_{j1})I(Z_{k2}\le Z_{i2}\wedge Z_{j2}) \\ &\quad - \left(\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} I(Z_{j1}\le Z_{i1}, Z_{j2}\le Z_{i2})\right)^2, \qquad (3.2.2)\end{aligned}$$

$$\begin{aligned}\hat{\sigma}_s^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_N(u\wedge v)dH_N^1(u_1)dH_N^2(u_2)dH_N^1(v_1)dH_N^2(v_2) \\ &\quad - \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_N(u_1,u_2)dH_N^1(u_1)dH_N^2(u_2)\right)^2 \qquad (3.2.3) \\ &= \frac{1}{N^5}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{i=1}^{N}\sum_{h=1}^{N}\left(\sum_{l=1}^{N} I(Z_{l1}\le Z_{i1}\wedge Z_{h1})I(Z_{l2}\le Z_{j2}\wedge Z_{k2})\right) \\ &\quad - \left(\frac{1}{N^3}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N} I(Z_{l1}\le Z_{i1})I(Z_{l2}\le Z_{j2})\right)^2. \qquad (3.2.4)\end{aligned}$$

Since H_N converges almost surely to F when H_0 is true, and is bounded by 1, we can argue as in the proof of Theorem 3.1.8 and then apply dominated convergence to prove that $\hat{\sigma}_k^2$ and $\hat{\sigma}_s^2$ converge almost surely to σ_k^2 and σ_s^2.

3.3 Conclusion

By Theorem 3.1.8, 3.1.6 and Slutsky's theorem, $\frac{K_{nm}}{\hat{\sigma}_k} \Rightarrow N(0,1)$, and $\frac{S_{nm}}{\hat{\sigma}_s} \Rightarrow N(0,1)$. For a given level of α, the approximate critical values are:

$$K_\alpha = z_\alpha \hat{\sigma}_k$$

$$S_\alpha = z_\alpha \hat{\sigma}_s$$

For testing using the Kendall statistic, we reject the null hypothesis when $K_{nm} \geq K_\alpha$. For testing using the Spearman statistic, we reject the null hypothesis when $S_{nm} \geq S_\alpha$. The corresponding p-values of these statistics are:

$$1 - \Phi(\frac{K_{nm}}{\hat{\sigma}_k})$$

$$1 - \Phi(\frac{S_{nm}}{\hat{\sigma}_s})$$

where Φ is standard normal distribution.

Chapter 4

Empirical Results

In this chapter, we carry out an empirical evaluation of the performance of the test statistics proposed in Chapters 2 and 3 for a test of $H_0 : F = G$ vs. $H_1 : F < G$, applying the various tests to the same data set in each case. The three test statistics for distributions on $\Re$ are compared in Section 4.1. Comparisons of the three tests for distributions on $\Re^2$ are made in Section 4.2. Our conclusions and recommendations are given in Chapter 5. All of the R code for carrying out the simulations, calculating the estimators, and producing the tables is given in Appendix A.

4.1 Empirical results in $\Re$

In this section we compare the Kolmogorov-Smirnov, Mann-Whitney-Wilcoxon and P-P plot test statistics for a test of $H_0 : F = G$ vs. $H_1 : F < G$ for distributions F and G on $\Re$. Our simulations are based on samples from uniform, normal and exponential distributions. In Table 4.1, we assume $F = G$ and compare different actual levels of the tests with the nominal (asymptotic) level for various samples sizes, n and m. In Table 4.2, we compare the power of the three statistics with samples from exponential distributions. In Table 4.3, we compare the power of the three statistics with samples from normal distributions. The R code used to calculate each of the estimators and then to produce the tables is given in Appendix A.1.

In Table 4.1, we compare the true values of different levels of tests with the nominal level for different combinations of sample sizes $50, 100, 200$. Column 'n' is the sample size from F. Column 'm' is the sample size from G. Column 'Alpha' is the true level of the tests. Column 'Repeats' is the number of times we repeat the tests for a fixed value of 'n' and 'm'. Column 'ksRejCount' is the number of rejections using the Kolmogorov-Smirnov statistic in 'Repeats' number of repetitions. Column 'ksRejProp' is the proportion of rejections in the total number of repetitions. Similarly Columns 'wxRejCount', 'wxRejProp' give the counts and proportions for the Wilcoxon test and 'ppRejCount', 'ppRejProp' for $P-P$ plots. Since all three test statistcs in $\Re$ are non-parametric, we can calculate these statistics using uniform samples for all possible distributions F and G in $\Re$. By comparing columns 'ksRejProp', 'wxRejProp' and 'ppRejProp', we notice that the Mann-Whitney-Wilcoxon estimation of α is almost always the closest to the true level.

Table 4.1: Level of tests comparison

TestID	n	m	Alpha	Repeat	ksRejCount	ksRejProp	wxRejCount	wxRejProp	ppRejCount	ppRejProp
1	50	50	0.05	2000	57	0.0285	90	0.045	66	0.033
2	50	50	0.05	6000	205	0.03416667	278	0.04633333	194	0.03233333
3	50	50	0.1	2000	170	0.085	209	0.1045	199	0.0995
4	50	50	0.1	6000	589	0.09816667	600	0.1	534	0.089
5	50	100	0.05	2000	91	0.0455	94	0.047	82	0.041
6	50	100	0.05	6000	236	0.03933333	299	0.04983333	208	0.03466667
7	50	100	0.1	2000	159	0.0795	177	0.0885	164	0.082
8	50	100	0.1	6000	474	0.079	574	0.09566667	529	0.08816667
9	50	200	0.05	2000	90	0.045	97	0.0485	91	0.0455
10	50	200	0.05	6000	265	0.04416667	303	0.0505	268	0.04466667
11	50	200	0.1	2000	205	0.1025	209	0.1045	186	0.093
12	50	200	0.1	6000	558	0.093	625	0.1041667	551	0.09183333
13	100	50	0.05	2000	87	0.0435	92	0.046	71	0.0355
14	100	50	0.05	6000	229	0.03816667	301	0.05016667	213	0.0355
15	100	50	0.1	2000	152	0.076	198	0.099	173	0.0865
16	100	50	0.1	6000	550	0.09166667	549	0.0915	500	0.08333333
17	100	100	0.05	2000	70	0.035	85	0.0425	66	0.033
18	100	100	0.05	6000	244	0.04066667	288	0.048	242	0.04033333
19	100	100	0.1	2000	142	0.071	189	0.0945	154	0.077
20	100	100	0.1	6000	434	0.07233333	614	0.1023333	475	0.07916667
21	100	200	0.05	2000	97	0.0485	94	0.047	104	0.052
22	100	200	0.05	6000	285	0.0475	263	0.04383333	269	0.04483333
23	100	200	0.1	2000	169	0.0845	207	0.1035	172	0.086
24	100	200	0.1	6000	553	0.09216667	606	0.101	483	0.0805
25	200	50	0.05	2000	79	0.0395	100	0.05	102	0.051
26	200	50	0.05	6000	262	0.04366667	318	0.053	294	0.049
27	200	50	0.1	2000	196	0.098	195	0.0975	190	0.095
28	200	50	0.1	6000	516	0.086	636	0.106	606	0.101
29	200	100	0.05	2000	98	0.049	91	0.0455	94	0.047
30	200	100	0.05	6000	265	0.04416667	292	0.04866667	271	0.04516667
31	200	100	0.1	2000	187	0.0935	190	0.095	149	0.0745
32	200	100	0.1	6000	544	0.09066667	595	0.09916667	527	0.08783333
33	200	200	0.05	2000	69	0.0345	110	0.055	103	0.0515
34	200	200	0.05	6000	247	0.04116667	302	0.05033333	254	0.04233333
35	200	200	0.1	2000	188	0.094	210	0.105	176	0.088
36	200	200	0.1	6000	529	0.08816667	575	0.09583333	525	0.0875

In Table 4.2, we did the power calculation using samples from exponential distributions $F(x) = 1 - e^{-x}$, $G(x) = 1 - e^{-2x}, x > 0$. We notice that the Mann-Whitney-Wilcoxon test is almost always more powerful than Kolmogorov and $P - P$ plot statistics.

Table 4.2: Power comparison from exponential

TestID	n	m	Alpha	Repeat	ksRejCount	ksRejProp	wxRejCount	wxRejProp	ppRejCount	ppRejProp
1	50	50	0.05	2000	1636	0.818	1811	0.9055	1636	0.818
2	50	50	0.05	6000	4955	0.8258333	5410	0.9016667	4955	0.8258333
3	50	50	0.1	2000	1847	0.9235	1916	0.958	1847	0.9235
4	50	50	0.1	6000	5544	0.924	5712	0.952	5544	0.924
5	50	100	0.05	2000	1853	0.9265	1937	0.9685	1853	0.9265
6	50	100	0.05	6000	5457	0.9095	5751	0.9585	5457	0.9095
7	50	100	0.1	2000	1937	0.9685	1960	0.98	1937	0.9685
8	50	100	0.1	6000	5818	0.9696667	5903	0.9838333	5818	0.9696667
9	50	200	0.05	2000	1924	0.962	1955	0.9775	1924	0.962
10	50	200	0.05	6000	5731	0.9551667	5847	0.9745	5731	0.9551667
11	50	200	0.1	2000	1963	0.9815	1981	0.9905	1963	0.9815
12	50	200	0.1	6000	5889	0.9815	5921	0.9868333	5889	0.9815
13	100	50	0.05	2000	1867	0.9335	1948	0.974	1867	0.9335
14	100	50	0.05	6000	5561	0.9268333	5807	0.9678333	5561	0.9268333
15	100	50	0.1	2000	1936	0.968	1962	0.981	1936	0.968
16	100	50	0.1	6000	5807	0.9678333	5906	0.9843333	5807	0.9678333
17	100	100	0.05	2000	1973	0.9865	1987	0.9935	1973	0.9865
18	100	100	0.05	6000	5905	0.9841667	5958	0.993	5905	0.9841667
19	100	100	0.1	2000	1987	0.9935	1999	0.9995	1987	0.9935
20	100	100	0.1	6000	5975	0.9958333	5989	0.9981667	5975	0.9958333
21	100	200	0.05	2000	1994	0.997	1998	0.999	1994	0.997
22	100	200	0.05	6000	5992	0.9986667	5994	0.999	5992	0.9986667
23	100	200	0.1	2000	1997	0.9985	2000	1	1997	0.9985
24	100	200	0.1	6000	5998	0.9996667	5999	0.9998333	5998	0.9996667
25	200	50	0.05	2000	1947	0.9735	1975	0.9875	1947	0.9735
26	200	50	0.05	6000	5800	0.9666667	5915	0.9858333	5800	0.9666667
27	200	50	0.1	2000	1974	0.987	1989	0.9945	1974	0.987
28	200	50	0.1	6000	5945	0.9908333	5972	0.9953333	5945	0.9908333
29	200	100	0.05	2000	1996	0.998	1998	0.999	1996	0.998
30	200	100	0.05	6000	5989	0.9981667	5994	0.999	5989	0.9981667
31	200	100	0.1	2000	2000	1	2000	1	2000	1
32	200	100	0.1	6000	5998	0.9996667	6000	1	5998	0.9996667
33	200	200	0.05	2000	2000	1	2000	1	2000	1
34	200	200	0.05	6000	6000	1	5999	0.9998333	6000	1
35	200	200	0.1	2000	2000	1	2000	1	2000	1
36	200	200	0.1	6000	6000	1	6000	1	6000	1

In Table 4.3, we did the power calculation using samples from normal distributions $F \stackrel{D}{=} N(0.5, 1)$, $G \stackrel{D}{=} N(0, 1)$. We notice that the Mann-Whitney-Wilcoxon test is almost always more powerful than Kolmogorov and $P - P$ plots statistics.

Table 4.3: Power comparison from normal

TestID	n	m	Alpha	Repeat	ksRejCount	ksRejProp	wxRejCount	wxRejProp	ppRejCount	ppRejProp
1	50	50	0.05	2000	1271	0.6355	1576	0.788	1271	0.6355
2	50	50	0.05	6000	3775	0.6291667	4711	0.7851667	3775	0.6291667
3	50	50	0.1	2000	1608	0.804	1746	0.873	1608	0.804
4	50	50	0.1	6000	4802	0.8003333	5248	0.8746667	4802	0.8003333
5	50	100	0.05	2000	1553	0.7765	1778	0.889	1553	0.7765
6	50	100	0.05	6000	4585	0.7641667	5271	0.8785	4585	0.7641667
7	50	100	0.1	2000	1764	0.882	1880	0.94	1764	0.882
8	50	100	0.1	6000	5226	0.871	5613	0.9355	5226	0.871
9	50	200	0.05	2000	1669	0.8345	1845	0.9225	1669	0.8345
10	50	200	0.05	6000	5096	0.8493333	5555	0.9258333	5096	0.8493333
11	50	200	0.1	2000	1845	0.9225	1923	0.9615	1845	0.9225
12	50	200	0.1	6000	5526	0.921	5772	0.962	5526	0.921
13	100	50	0.05	2000	1482	0.741	1729	0.8645	1482	0.741
14	100	50	0.05	6000	4565	0.7608333	5270	0.8783333	4565	0.7608333
15	100	50	0.1	2000	1742	0.871	1868	0.934	1742	0.871
16	100	50	0.1	6000	5256	0.876	5617	0.9361667	5256	0.876
17	100	100	0.05	2000	1815	0.9075	1925	0.9625	1815	0.9075
18	100	100	0.05	6000	5427	0.9045	5796	0.966	5427	0.9045
19	100	100	0.1	2000	1921	0.9605	1969	0.9845	1921	0.9605
20	100	100	0.1	6000	5711	0.9518333	5903	0.9838333	5711	0.9518333
21	100	200	0.05	2000	1933	0.9665	1976	0.988	1933	0.9665
22	100	200	0.05	6000	5823	0.9705	5943	0.9905	5823	0.9705
23	100	200	0.1	2000	1976	0.988	1995	0.9975	1976	0.988
24	100	200	0.1	6000	5918	0.9863333	5983	0.9971667	5918	0.9863333
25	200	50	0.05	2000	1694	0.847	1842	0.921	1694	0.847
26	200	50	0.05	6000	5087	0.8478333	5562	0.927	5087	0.8478333
27	200	50	0.1	2000	1859	0.9295	1920	0.96	1859	0.9295
28	200	50	0.1	6000	5526	0.921	5781	0.9635	5526	0.921
29	200	100	0.05	2000	1941	0.9705	1983	0.9915	1941	0.9705
30	200	100	0.05	6000	5828	0.9713333	5947	0.9911667	5828	0.9713333
31	200	100	0.1	2000	1965	0.9825	1990	0.995	1965	0.9825
32	200	100	0.1	6000	5930	0.9883333	5989	0.9981667	5930	0.9883333
33	200	200	0.05	2000	1993	0.9965	2000	1	1993	0.9965
34	200	200	0.05	6000	5971	0.9951667	5994	0.999	5971	0.9951667
35	200	200	0.1	2000	1997	0.9985	2000	1	1997	0.9985
36	200	200	0.1	6000	5992	0.9986667	6000	1	5992	0.9986667

Conclusion: Kolmogorov-Smirnov and $P - P$ plot statistics results are very similar for level of tests and are identical for powers. Mann-Whitney-Wilcoxon performs best both in level of test and power.

4.2 Empirical results in $\Re^2$

In this section we compare the marginal Mann-Whitney-Wilcoxon, Spearman and Kendall test statistics for a test of $H_0 : F = G$ vs. $H_1 : F < G$ for distributions F and G on $\Re^2$. We examine various scenarios. In section 4.2.1, we consider an alternative in which G is a shifted version of F. In this situation, we note that F and G have identical copulas but different marginals. In the next two sections, we look at alternatives in which the shift assumption is violated: in section 4.2.2, we consider the case where F and G have equal marginal distributions but different copula structures and in section 4.2.3 we allow both the marginal distributions and copula structures to differ.

The algorithms and R-code used to calculate each of the three estimators is given in Appendix A.2. The calculation of the Kendall statistic is particularly time consuming, and it was necessary to find an efficient algorithm which is described in detail in Appendix A.2.3. Because of the length of time needed to both carry out the simulations and then produce the statistics, the tests are carried out on five simulated data sets and in each case, the computed p-values are displayed.

4.2.1 Shifted distributions

In this section, we compare the Mann-Whitney-Wilcoxon, Kendall and Spearman statistics using multivariate normal samples. All code is given in Appendix A.3.1. Let F and G be bivariate normal distributions; we compare the marginal Mann-

Whitney-Wilcoxon, Kendall and Spearman statistics for different sample sizes under $H_0 : F = G$ and $H_1 : F < G$.

In Table 4.4, we compare the three test statistics for sample sizes of $n = m = 50$ under $H_0 : F = G$ for five samples where $F, G \sim N(\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 10 & 3 \\ 3 & 2 \end{smallmatrix}\right))$. In Table 4.5, we compare the three test statistics for sample sizes of $n = m = 50$ when $F \sim N(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 10 & 3 \\ 3 & 2 \end{smallmatrix}\right))$ and $G \sim N(\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 10 & 3 \\ 3 & 2 \end{smallmatrix}\right))$. The 'Test ID' column is the number of the test. 'MWW p-value', 'Kendall p-value' and 'Spearman p-value' are the corresponding p-values. We notice that when H_0 is true and the sample sizes are small, none of the statistics seem to out-perform the others. When H_1 is true, mostly the p-value of the Kendall statistic is less than Mann-Whitney-Wilcoxon's and Spearman's but not considerably. However, when the sample sizes n, m increase to 200, Table 4.6, 4.7 show that under H_0, the three statistics' p-values have no significant differences, and under H_1, the Mann-Whitney-Wilcoxon p-value is almost always the smallest among the three, but all three statistics' p-values are very small.

Therefore, we can conclude that if the shifted distribution condition is true and $\Delta_1 < 0, \Delta_2 < 0$, the marginal Mann-Whitney-Wilcoxon statistic should be chosen for distribution comparisons.

Table 4.4: $n = m = 50$ under H_0

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.7763723	0.4425461	0.4455034
2	0.1959824	0.5	0.5495664
3	0.4200402	0.101315	0.1218223
4	0.1413844	0.05203059	0.1125096
5	0.2956415	0.7267877	0.6976813

Table 4.5: $n = m = 50$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.0001840377	0.0001689324	0.003710840
2	0.002211255	0.0003903628	0.002457298
3	0.03252237	0.001647246	0.006414776
4	4.754236e-05	2.943365e-07	1.450525e-05
5	0.05509695	0.01097064	0.02056974

Table 4.6: $n = m = 200$ under H_0

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.3238428	0.1002369	0.1700506
2	0.1889213	0.07346214	0.08200816
3	0.1346824	0.3589066	0.4246917
4	0.4645585	0.2287688	0.3115745
5	0.7481738	0.9009968	0.8641205

Table 4.7: $n = m = 200$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	2.772850e-16	4.326019e-11	1.214720e-08
2	4.480498e-10	1.260663e-10	1.516377e-08
3	3.448109e-15	7.099627e-13	2.014087e-10
4	1.560869e-09	2.408458e-07	6.204493e-06
5	1.03643e-10	1.307473e-09	1.598853e-07

4.2.2 Equal marginals, different copulas

In this section we will compare the three statistics when F and G have equal marginals but different copulas. Without loss of generality, the marginals are assumed to be $U(0,1)$ and the copulas are FGM copula with different parameters.

Definition 4.2.1 *The FGM copula (Farlie-Gumbel-Morgenstern; ref. [2]) is defined as:*

$$C_\theta(u,v) = uv + \theta uv(1-u)(1-v) \quad where (u,v) \in [0,1]^2$$

Appendix A.2.1 describes the algorithm and the R code to simulates the FGM copula in detail and we will use it to produce the samples for the three statistics.

Tests under H_1

In the scenario that F and G have the same marginal distribution but differnt FGM copulas, we assume that $H_1 : F < G$ is true. Suppose $X_1, \ldots X_n$ and $Y_1, \ldots Y_m$ are samples from F and G with same marginal distributions $F_1 = F_2 = G_1 = G_2 \stackrel{D}{=} U(0,1)$ and different FGM copulas $F(x_1,x_2) = C_{\theta_X}(F_1(x_1), F_2(x_2))$, $G(y_1,y_2) = C_{\theta_Y}(G_1(y_1), G_2(y_2))$ where $\theta_X < \theta_Y$. The R-Code in Appendix A.3.2 is used to compare the performance of the three test statistics with different values of θ_X, θ_Y and we present the results in the following tables.

In Tables 4.8 and 4.9 we compare the three statistics' p-values when $\theta_X = -1, \theta_Y = -0.5$ for $n = m = 50$ and $n = m = 200$ respectively. In Tables 4.10 and 4.11 we compare the three statistics' p-values when $\theta_X = -1, \theta_Y = 0.5$ for $n = m = 50$ and $n = m = 200$ respectively. In Tables 4.12 and 4.13 we compare the three statistics' p-values when $\theta_X = -1, \theta_Y = 1$ for $n = m = 50$ and $n = m = 200$

respectively.

Table 4.8 and 4.9 show that when sample sizes are small the three statistics have no significant differences and won't reject the null hypothesis for small values of α. When the sample sizes increase to 200, all three statistics' p-values decrease and may lead to rejection of the null hypothesis, but Mann-Whitney-Wilcoxon is still unlikely to reject. Moreover, the Kendall statistic has the smallest p-value. The Spearman statistic has the middle level p-value but it still performs well and its p-value is close to that of the Kendall statistic.

Table 4.8: $n = m = 50$, $\theta_X = -1, \theta_Y = -0.5$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.3238189	0.8502385	0.8252522
2	0.3562291	0.4496801	0.4435727
3	0.923675	0.4092942	0.3937101
4	0.7281376	0.6620396	0.6564226
5	0.3640727	0.8755698	0.8100370

Table 4.9: $n = m = 200$, $\theta_X = -1, \theta_Y = -0.5$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.9633662	0.0530085	0.1020852
2	0.8190913	0.2431211	0.2793360
3	0.1144726	0.0003305783	0.001552441
4	0.5683541	0.02232232	0.03486748
5	0.06752812	0.007547048	0.02036734

Table 4.10 indicates when the FGM copulas differ more ($\theta_X = -1, \theta_Y = 0.5$) and $n = m = 50$, the p-value of Kendall statistic decreases and is more likely to lead to rejection of the null hypothesis, while the Mann-Whitney-Wilcoxon's p-value remains large. The Spearman statistic's p-value also decreases and remains close to Kendall statistic's p-value.

Table 4.10: $n = m = 50$, $\theta_X = -1, \theta_Y = 0.5$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.9212954	0.2913416	0.3413597
2	0.8864444	0.04423535	0.0722065
3	0.5060633	0.00378754	0.01040310
4	0.3177963	0.1771308	0.2218798
5	0.6483741	0.05405529	0.08249608

In Tabel 4.11, $n = m = 200$, $\theta_X = -1, \theta_Y = 0.5$, we got more small p-values for Kendall and Spearman statistics. Kendall statistic still has the best performance, and Mann-Whitney-Wilcoxon is unlikely to reject the null hypothesis.

Table 4.11: $n = m = 200$, $\theta_X = -1, \theta_Y = 0.5$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.7783239	0.000523714	0.001946464
2	0.2008122	0.1695921	0.2038195
3	0.7017442	0.01151214	0.02372157
4	0.9728018	0.002611836	0.006678975
5	0.7696677	0.03210502	0.05999754

In Table 4.12, $n = m = 50$, $\theta_X = -1, \theta_Y = 1$, the two FGM copula differ more and the average result of all the p-values does not change significantly. Kendall statistic still have the smallest p-value and Spearman statistic's p-value remains to close to Kendall statistic's p-value.

Table 4.12: $n = m = 50$, $\theta_X = -1, \theta_Y = 1$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.3400971	0.2946489	0.3574400
2	0.6990412	0.04939621	0.0794407
3	0.6592368	0.1108608	0.1536002
4	0.4450927	0.004121707	0.009913544
5	0.5241751	0.02681852	0.0426612

Table 4.13 shows that when sample sizes are large and the FGM copula differ a lot, Mann-Whitney-Wilcoxon statistic is still unlikely to reject the null hypothesis, while Kendall and Spearman statistics are both likely to reject the null hypothesis.

Table 4.13: $n = m = 200$, $\theta_X = -1, \theta_Y = 1$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.551322	5.316721e-05	0.0003826404
2	0.5515656	0.01395351	0.02693787
3	0.3119021	0.04906331	0.06973668
4	0.1699126	2.824949e-06	3.597677e-05
5	0.2652574	0.1227713	0.1436605

Conclusion: When F and G have same the marginal distributions but different copulas, the Kendall statistic performs best, followed closely by Spearman statistic. As expected, since the Mann-Whitney-Wilcoxon statistic compares only marginal distributions and the shift assumption is violated, this test does not perform well.

4.2.3 Different marginals, different copulas

In this section, we will compare the three test statistics when F and G have different exponential marginals and different FGM copulas. In this case, we expect that the three statistics will all produce small p-values. Moreover, we will let $F_1 = F_2 = G_1 \sim exp(1)$ and make G_2 closer and closer to F_2 and see how the three statistics perform.

By Sklar's theorem, any bivariate distribution can be written in form of some copula, for example, $F(x, y) = C(F_1(x), F_2(y))$. Hence, given F and C we could simulate observations (U, V) from C first and then use inversion to simulate $(X_1, X_2) = (F_1^{-1}(U), F_2^{-1}(V))$ from F.

Suppose F and G have FGM copulas with parameter $\theta_X = -1$, $\theta_Y = 1$, $F_1 = F_2 \sim exp(1)$, $G_1 \sim exp(1)$, and $G_2 \sim exp(1.125)$, $exp(1.25)$, $exp(1.5)$ or $exp(1.75)$. We will compare the three test statistics with sample size $n = m = 200$ in these scenarios. The code that produces the following tables may be found in Appendix A.3.3.

The scenario in Table 4.14 is when G_2 is very different from F_2. All three test statistics reject the null hypothesis even for very small levels of significance. The Kendall statistic has the smallest p-value, Spearman statistic has the second smallest p-value, and Mann-Whitney-Wilcoxon has the largest p-value.

In Table 4.15, G_2 get a little step closer to F_2. The Mann-Whitney-Wilcoxon's p-value increases but may still reject H_0 when α is not too small. The Kendall and Spearman statistics' p-values remain very small and will lead to rejection of the null hypothesis.

In Table 4.16, G_2 gets even closer to F_2. The Mann-Whitney-Wilcoxon's p-value gets larger and is more likely to fail to reject the null hypothesis, while the Kendall

Table 4.14: $n = m = 200$, $G_2 \sim exp(1.75)$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	1.381224e-06	1.925739e-10	4.502919e-08
2	6.585361e-09	9.294508e-12	3.861385e-09
3	4.385588e-05	9.628231e-08	2.828138e-06
4	0.005641764	6.165296e-09	5.821362e-07
5	2.502747e-07	9.712652e-10	1.361892e-07

Table 4.15: $n = m = 200$, $G_2 \sim exp(1.5)$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.000691196	1.296718e-08	8.349835e-07
2	0.0008037985	7.738197e-08	3.362226e-06
3	3.321108e-06	1.678560e-09	1.783553e-07
4	0.007022952	8.579551e-06	8.621545e-05
5	0.1693699	1.118473e-06	2.639814e-05

and Spearman statistics' p-values still remain very small.

Conclusion: The Kendall statistic performs the best among the three statistics and it should be chosen in general. The Spearman statistic also performs well, its p-value is close to Kendall statistic's p-value But the advantage of Spearman statistic is that when the sample size is large, its calculation time is much less than Kendall. The Mann-Whitney-Wilcoxon statistic may fail when the marginals of F and G do not differ too much, and it should not be chosen when the shift condition is uncertain.

Table 4.16: $n = m = 200$, $G_2 \sim exp(1.25)$ under H_1

Test ID	MWW p-value	Kendall p-value	Spearman p-value
1	0.01563369	5.594523e-05	0.0004390573
2	0.1640488	1.728735e-05	0.0001543052
3	0.006467424	2.905821e-09	1.853460e-07
4	0.08123819	1.819638e-06	2.581783e-05
5	0.9834702	0.001161249	0.004108039

Chapter 5

Concluding Remarks

Summary of main contributions of the thesis:

- Definition of two new bivariate analogues of the Mann-Whitney-Wilcoxon statistic
- Analysis of asymptotic distributions of both statistics
- Detailed empirical comparison of the performance of all three statistics under various scenarios

Recommendation: When testing $H_0 : F = G$ vs. $H_1 : F < G$, if the sample size is small, the Kendall statistic should be chosen as the candidate. Even when the shift condition is satisfied, the Kendall statistic also performs very well. If the sample size is very large, the Spearman statistic should be chosen, since it does not require the shift condition and it performs better than the marginal Mann-Whitney-Wilcoxon statistic in general. Moreover, the Spearman statistic consumes much less computation time compared to the Kendall statistic, and its performance is close to that of the Kendall statistic.

Appendix A

R-Codes

A.1 R-code for Section 4.1

A.1.1 R-code for Table 4.1

```
######################################
# P-P-plot statistic in R
# returns a list which consists of
# p-value and the statistic value.
######################################
PPlotTest<-function(X, Y)
{
    n = length(X);
    m = length(Y);
    parts = c(rep(0,n));
    xOrdered = sort(X);
    for (k in 1:n)
    {
        # k-1/n < p <= k/n
```

```
        #F_n^-(p) and G_m(F_n-) is constant
        t = 0;
        for (j in 1:m)
        {
            if (Y[j] <= xOrdered[k])
            {
                t = t+1;
            }
        }
        # sup value for k-1/n < p <= k/n
        t = t/m - (k-1)/n;
        parts[k] = t;
    }
    stat = max(parts);
    pvalue = exp(-2*(sqrt(n*m/(n+m))*stat)^2);
    return(list(p.value = pvalue, statistic = stat));
}
#########################################
# One dimensional test of the three statistics
# n and m are sample size of F,G distributions
# alp is the level of the tests
# rep is the number of repitions you want it to do.
#########################################
OneDimTest<-function(n, m, alp, rep)
{
    ksRejects = 0;
    wxRejects = 0;
    ppRejects = 0;
```

```
for (i in 1:rep)
{
    X = runif(n);
    Y = runif(m);
    # kolmogorov-Smirnov test
    ksResult = ks.test(X,Y, alternative = "greater");
    #print(ksResult);
    if (ksResult$p.value < alp)
    {
        ksRejects = ksRejects + 1;
    }
    wxResult = wilcox.test(X,Y, alternative = "greater");
    #print(wxResult);
    if (wxResult$p.value < alp)
    {
        wxRejects = wxRejects + 1;
    }
    ppResult = PPlotTest(X,Y);
    #print(ppResult);
    if (ppResult$p.value < alp)
    {
        ppRejects = ppRejects + 1;
    }
}

return(list(ksRejCount = ksRejects, ksRejProp = ksRejects/rep,
            wxRejCount = wxRejects, wxRejProp = wxRejects/rep,
            ppRejCount = ppRejects, ppRejProp = ppRejects/rep));
```

```
}

##################################
# Test all kind of combinations
##################################
nm = c(50,100,200);

alp = c(0.05,0.10);

rep = c(2000,6000);

id = 1;

for (i in 1:3)
{
    for (j in 1:3)
    {
        for (k in 1:2)
        {
            for (l in 1:2)
            {
                tInfo = list(TestID = id, n = nm[i],
                             m = nm[j], Alpha = alp[k],
                             Repeat = rep[l]);
                id = id + 1;
                print(tInfo);
                result = OneDimTest(nm[i], nm[j], alp[k], rep[l]);
```

```
                print("*****OUTPUT*****");
                print(result);
                print("========================");
            }
        }
    }
}
```

A.1.2 R-code for Table 4.2, 4.3

```
############################################################
# Modified OneDimTest with additional options
# n and m are sample size of F,G distributions
# alp is the level of the tests.
# rep is the number of repetitions you want it to do.
# XGen is the X random sample generator function,
# for exmpale, rnorm, rexp.
# YGen is the Y random sample generator
# XArg is a parameter for XGen, if XGen = rnorm, then XArg is the mean
# of the normal distribution. If XGen = rexp, then XArg is the rate
# of exponential distribution. Similarly for YArg.
############################################################
OneDimTest<-function(n, m, alp, rep, XGen, YGen, XArg, YArg)
{
   ksRejects = 0;
   wxRejects = 0;
   ppRejects = 0;
   for (i in 1:rep)
```

```
{
    X = XGen(n,XArg);
    Y = YGen(m,YArg);
    # kolmogorov-Smirnov test
    ksResult = ks.test(X,Y, alternative = "less");
    #print(ksResult);
    if (ksResult$p.value < alp)
    {
        ksRejects = ksRejects + 1;
    }
    wxResult = wilcox.test(X,Y, alternative = "greater");
    #print(wxResult);
    if (wxResult$p.value < alp)
    {
        wxRejects = wxRejects + 1;
    }
    ppResult = PPlotTest(X,Y);
    #print(ppResult);
    if (ppResult$p.value < alp)
    {
        ppRejects = ppRejects + 1;
    }
}
return(list(ksRejCount = ksRejects, ksRejProp = ksRejects/rep,
 wxRejCount = wxRejects, wxRejProp = wxRejects/rep,
 ppRejCount = ppRejects, ppRejProp = ppRejects/rep));

}
```

```
nm = c(50,100,200);
alp = c(0.05,0.10);
rep = c(2000,6000);
dis = c(rexp, rnorm);
disName = c("Exp-Dist", "Norm-Dist");
id = 1;

for (s in 1:2)
{
   for (i in 1:3)
   {
       for (j in 1:3)
       {
           for (k in 1:2)
           {
               for (l in 1:2)
               {

                       tInfo = list(TestID = id, n = nm[i], m = nm[j],
                        Alpha = alp[k], Repeat = rep[l]);
                       id = id + 1;
                       print(tInfo);
                       result = NULL;
                       if (s == 1)
                       {
```

```
                            result = OneDimTest(nm[i], nm[j], alp[k],
                             rep[l], dis[[s]], dis[[s]], 1, 2);
                        }
                        else
                        {
                            result = OneDimTest(nm[i], nm[j], alp[k],
                             rep[l], dis[[s]], dis[[s]], 0.5, 0);
                        }

                        print("*****OUTPUT*****");
                        print(disName[s]);
                        print(result);
                        print("=======================================");

                }

            }
        }
    }
}
```

A.2 R-Code for Simulation and Statistic Computation

In this section, we will present the procedures and R-codes to simulate FGM copula samples and algorithms that compute the marginal Mann-Whitney-Wilcoxon, Kendall and Spearman statistics. All the codes are verified with trusted data sets or true values of the estimates. To use the R-codes, you will need to use R software and load the package 'MASS'. These codes presented in this section are the prerequisites to perform the codes in section A.3.

A.2.1 R-code for FGM Copula Simulation

According to Sklar's theorem, for every bivariate distribution $F(x_1, x_2)$ with marginal distributions F_1, F_2, there exist a copula C such that $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$, hence we can simulate F distributed samples by simulating the copula corresponding to F and use inversion or rejection method to simulate the samples $X_i = (X_{i,1}, X_{i,2})$, $i = 1, \ldots, n$.

we will use a particular type of copula, Farlie-Gumbel-Morgenstern, for our simulations of bivariate statistics. The Farlie-Gumbel-Morgenstern(FGM)(ref. [2]) copula is defined as:

$$C_\theta(u, v) = uv + \theta uv(1-u)(1-v). \tag{A.2.1}$$

The density function is:

$$f_\theta(u, v) = \frac{\partial C_\theta(u, v)}{\partial v \partial u} = 1 + \theta(1-2u)(1-2v). \tag{A.2.2}$$

The marginal densities of u and v are the uniform density on $[0, 1]$: $f_u(u) = f_v(v) = 1$, hence the conditional density of v given u is $f_{v|u}(v|u) = \frac{f_\theta(u,v)}{f_u(u)} = f_\theta(u, v)$. The conditional cdf of $V|U = u$ is:

$$F_{V|U}(v|u) = \int_0^v f_\theta(u,t)dt = \frac{\partial C_\theta(u,v)}{\partial u}. \tag{A.2.3}$$

$$\frac{\partial C_\theta(u,v)}{\partial u} = \theta(2u-1)v^2 + (1+\theta-2\theta u)v. \tag{A.2.4}$$

For fixed values of u and θ, the conditional cdf $F_{V|U}(v|u)$ is a univariate distribution function of v. By the inversion method, we can simulate the random variable V by the following steps:

1. Generate an observation u of $U \sim U[0,1]$
2. Generate an observation c of $C \sim U[0,1]$
3. Use u and applied to c to generate an observation v of $V \sim F_{V|U}(v|u)$. v must satisfies:

$$\underbrace{\theta(2u-1)}_{a}v^2 + \underbrace{(1+\theta-2\theta u)}_{b}v \underbrace{-c}_{c'} = 0 \tag{A.2.5}$$

Hence $v = \frac{-b\pm\sqrt{b^2-4ac'}}{2a}$. Since $F_{V|U}(v|u)$ is a non-discreasing function, there is at most a unique root inside $[0,1]$.

```
# n: number of sample to generate
# theta is the parameter in C_theta(u,v) = uv+theta*uv(1-u)(1-v)
#F_{V |U}(v|u) = theta*(2u-1)v^2+(1+theta-2*theta*u)v
# The return value is the a matrix of a form like this:
#   u1 u2 u3 u4 .u_n
#   v1 v2 v3  v4 .v_n
FGMSimulation<-function(n, theta)
{
```

```
ret = matrix(,2,n);
cnt = 0;
while (cnt<n)
{
  #Generate U
  u=runif(1);
  #Generate conditional cdf of V given u
  c=runif(1);
  b = 1+ theta-2*theta*u;
  a = theta*(2*u-1);
  delta = b^2+4*a*(c);
  if (delta < 0)
  {
    next;
  }
  v1 = (-b+sqrt(delta))/(2*a);
  v2 = (-b-sqrt(delta))/(2*a);
  if (v1>=0 && v1<=1)
  {
    cnt = cnt + 1;
    ret[1,cnt] = u;
    ret[2,cnt] = v1;
    next;
  }
  if (v2>=0 && v2<=1)
  {
    cnt = cnt + 1;
    ret[1,cnt] = u;
```

```
      ret[2,cnt] = v2;
      next;
    }
  }
  return(ret);
}
```

A.2.2 R-code for the marginal Mann-Whitney-Wilcoxon statistic

This function computes the marginal Mann-Whitney-Wilcoxon statistic. The parameter Z is the combined sample of X and Y with size n and m. p is the dimension of X and Y, which is always 2 in our case.

```
MVManWilTest<- function(z, p, n, m)
{
      stat = 0;
      R = matrix(, 0, n+m);
      for ( i in 1:p)
      {
            R = rbind(R,rank(z[i,]))
      }
      U = matrix(rep(0,p),p,1);
      for ( i in 1:p)
      {
            U[i,1] = sum(R[i,n+1:m])/(n+m+1)- m/2;
      }
      V = matrix(,p,p);
      N = n+m;
```

```
        for ( i in 1:p)
        {
              for (j in 1:p)
              {
                    V[i,j] = (m*n/(N^2*(N-1)*(N+1)^2))*
                    (crossprod(R[i,],R[j,]) - (1/4)*N*(N+1)^2)
              }
        }
        stat = t(U) %*% ginv(N*V) %*% U;
        return (stat);
}
```

A.2.3 R-code for Kendall statistic

In order to reduce the time of computing the Kendall statistic. We propose a very efficient algorithm to dramatically reduce the computation time. The following algorithm aims to reduce the time to compute the estimated Kendall variance, which is in the function 'EstimateSigmaSqOpt'.

Algorithm:

1. let $Rnk_{s,t}$ be the rank of $X_{s,t}$ where $s = 1, 2, t = 1...n$
2. Sort Rnk by its first components(i.e. sort $Rnk_{1,\cdot}$). When sorting one of the components, keep the other component together with the sorting component.
3. Let $Rnk2$ be a copy of Rnk, sort $Rnk2$ by its second component.
4. let $h1$ be $Rnk_{1,i}$, $h2 = min(Rnk_{2,i}, Rnk_{2,j})$ $i \leq j$
5. if $h1 \leq h2$ check the number of columns in matrix Rnk such that $Rnk_{2,k} \leq h2$ for the first $h1$ number of columns.

6. else check the number of columns in matrix $Rnk2$ such that $Rnk_{1,k} \leq h1$ for the first $h2$ number of columns.

7. Continue and add up all the countings in 4,5,6

8. divide the counting by n^3

In the following code, the bubble sort algorithm has been used. However, the speed can be further optimized if we choose the quick sort algorithm, since the expected performance of quick sort is $O(nlogn)$, while the bubble sort is $O(n^2)$. We did not use the quick sort algorithm, since the number of nested functions allowed in R is small, and we could encounter problems with large samples.

```
#EmpIntegral(Y,X) =\int\intG_mdF_n
EmpIntegral<-function(sample1, sample2)
{
   m = dim(sample1)[2]; # size of Y sample
   n = dim(sample2)[2]; # size of X sample
   ret = 0;
   for (i in 1:n)
   {
      for (j in 1:m)
      {
         if (sample1[1,j] <= sample2[1,i] && sample1[2,j] <= sample2[2,i])
         {
            ret = ret + 1/(n*m);
         }
      }


   }
```

```
   return(ret);
}

# the first parameter mtx is the matrix of the form
# Z_11, Z_12,...Z_1n
# Z_21, Z_22,...Z_2n
# The second parameter snd means, if snd = TRUE,
# then we sort mtx based on its first component,
# otherwise, sort its second component.
bubbleSort<-function(mtx, snd = FALSE)
{
  if (snd == TRUE)
  {
    t=mtx[1,];
    mtx[1,] = mtx[2,];
    mtx[2,] = t;
  }
  n = dim(mtx)[2];
  i = n;

  while (i > 1)
  {
    for (j in 2:i)
    {
      if (mtx[1,j-1] > mtx[1,j])
      {
        t1 = mtx[1,j-1];
        t2 = mtx[2,j-1];
```

```
        mtx[1,j-1] = mtx[1,j];
        mtx[2,j-1] = mtx[2,j];
        mtx[1,j] = t1;
        mtx[2,j] = t2;
      }
    }
    i = i - 1;
  }

  if (snd == TRUE)
  {
    t=mtx[1,];
    mtx[1,] = mtx[2,];
    mtx[2,] = t;
  }
  return(mtx);
}


# Estimation of \sigma^2_F
# Here we use the algorithm proposed above to
# calculate the estimated variance
EstimateSigmaSqOpt<-function(X)
{
   n = dim(X)[2];
   Rnks = matrix(,2,n);
   Rnks[1,] = rank(X[1,]);
   Rnks[2,] = rank(X[2,]);
```

```
Rnks2 = matrix(Rnks, 2, n, FALSE);
Rnks = bubbleSort(Rnks);
Rnks2 = bubbleSort(Rnks2, TRUE);
ret = 0;
for(i in 1:n)
{
   for (j in i:n)
   {
      cnt = 0;
      h1 = Rnks[1,i];
      h2 = min(Rnks[2,i], Rnks[2,j]);
         if (h1 <= h2)
         {
            for (k in 1:h1)
            {
                if (Rnks[2,k] <= h2)
                {
                   cnt = cnt + 1;
                }
            }
         }
         else
         {
            for (k in 1:h2)
            {
                if (Rnks2[1,k] <= h1)
                {
                   cnt = cnt + 1;
```

```
                    }
                }
            }
         if (i != j)
         {
            cnt = cnt*2;
         }
         ret = ret + cnt;
      }
   }
   ret = ret/n^3;
   #print(ret);
   #print((EmpIntegral(X,X))^2);
   ret = ret -  (EmpIntegral(X,X))^2; # Equation 35
   return(ret);
}


# H_{n+m} is the empirical function created by
# combining X and Y sample,
# which is denoted by Z.
# This is true under H0 hypothesis
KendallStat<-function(X,Y)
{
   Z = cbind(X,Y);
   n = dim(Z)[2];
   m = dim(Y)[2];
   Wnm = sqrt((n*m)/(n+m)) *  (EmpIntegral(Y,Z) - EmpIntegral(X,Z));
```

```
sig =  sqrt(EstimateSigmaSqOpt(Z));
stat = Wnm/sig;
pval = pnorm(stat, 0, 1, lower=FALSE);
return(c(stat,pval));

}
```

A.2.4 R-code for Spearman statistic

In the equation (2.2.8), the summation can be rewritten as follows:

$$
\begin{aligned}
S_{nm} &= \sqrt{\frac{nm}{n+m}}\left[\frac{1}{m}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{h=1}^{m} I(Y_{h1} \le Z_{i1}) I(Y_{h2} \le Z_{j2})\right] \\
&\quad -\sqrt{\frac{nm}{n+m}}\left[\frac{1}{n}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{h=1}^{n} I(X_{h1} \le Z_{i1}) I(X_{h2} \le Z_{j2})\right] \\
&= \sqrt{\frac{nm}{n+m}}\left[\frac{1}{m}\frac{1}{N^2}\sum_{h=1}^{m}\left(\sum_{i=1}^{N} I(Y_{h1} \le Z_{i1})\right)\left(\sum_{j=1}^{N} I(Y_{h2} \le Z_{j2})\right)\right] \\
&\quad -\sqrt{\frac{nm}{n+m}}\left[\frac{1}{n}\frac{1}{N^2}\sum_{h=1}^{n}\left(\sum_{i=1}^{N} I(X_{h1} \le Z_{i1})\right)\left(\sum_{j=1}^{N} I(X_{h2} \le Z_{j2})\right)\right] \\
&= \sqrt{\frac{nm}{n+m}}\left[\frac{1}{m}\frac{1}{N^2}\sum_{h=1}^{n}(N - R_{Y_{h1}} + 1)(N - R_{Y_{h2}} + 1)\right] \\
&\quad -\sqrt{\frac{nm}{n+m}}\left[\frac{1}{n}\frac{1}{N^2}\sum_{h=1}^{n}(N - R_{X_{h1}} + 1)(N - R_{X_{h2}} + 1)\right] \qquad \text{(A.2.6)}
\end{aligned}
$$

where $R_{Y_{h1}}$ is the rank of Y_{h1} in the combined sample in the first component. $R_{Y_{h2}}$, $R_{X_{h1}}$ and $R_{X_{h2}}$ are defined similarly.

Morever, we can also simplify the estimator of the variance of Spearman statistic. In the equation (3.2.3), both the first part and second part can be written as follows:

- **First part:**

$$
\begin{aligned}
&\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{i=1}^{N}\sum_{h=1}^{N}\left(\sum_{l=1}^{N} I(Z_{l1} \le Z_{i1} \wedge Z_{h1}) I(Z_{l2} \le Z_{j2} \wedge Z_{k2})\right) \\
=\ & \sum_{l=1}^{N}\left(\sum_{i=1}^{N}\sum_{h=1}^{N} I(Z_{l1} \le Z_{i1} \wedge Z_{h1})\right)\left(\sum_{j=1}^{N}\sum_{k=1}^{N} I(Z_{l2} \le Z_{j2} \wedge Z_{k2})\right) \\
=\ & \sum_{l=1}^{N}\left(\sum_{i=R_{Z_{l1}}}^{N}\ \sum_{h=R_{Z_{l1}}}^{N} I(Z_{l1} \le Z_{i1} \wedge Z_{h1})\right)\left(\sum_{j=R_{Z_{l2}}}^{N}\ \sum_{k=R_{Z_{l2}}}^{N} I(Z_{l2} \le Z_{j2} \wedge Z_{k2})\right) \\
=\ & \sum_{l=1}^{N}(N - R_{Z_{l1}} + 1)^2 (N - R_{Z_{l2}} + 1)^2 \qquad \text{(A.2.7)}
\end{aligned}
$$

- **Second part:**

$$
\begin{aligned}
&\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N} I(Z_{l1} \le Z_{i1}) I(Z_{l2} \le Z_{j2}) \\
=\ & \sum_{l=1}^{N}\left(\sum_{i=1}^{N} I(Z_{l1} \le Z_{i1})\right)\left(\sum_{j=1}^{N} I(Z_{l2} \le Z_{j2})\right) \\
=\ & \sum_{l=1}^{N}(N - R_{Z_{l1}} + 1)(N - R_{Z_{l2}} + 1) \qquad \text{(A.2.8)}
\end{aligned}
$$

The performence of the following code is very efficient. For X and Y with sample size 600 and 600, the test only takes around 30 seconds, which is much better than using the Kendall statistic. The reason is that, in the Kendall statistic, even though we have fewer summations, the summation indicators are dependent, while in Spearman's method, the summations are independent which reduces the computation time.

Notice that the following code will use the function bubbleSort which is declared in Appendix A.2.3

```
##########################################################
# Bubble sort.
```

```
# Optimize the EmpIntegral Speed for Spearman statistic
# If the parameter snd is TRUE, then we will sort the
# second component in the sample vectors instead of
# the first component.
#########################################################
bubbleSortSpearman<-function(mtx, snd = FALSE)
{
  if (snd == TRUE)
  {
    t=mtx[1,];
    mtx[1,] = mtx[2,];
    mtx[2,] = t;
  }
  n = dim(mtx)[2];
  i = n;

  while (i > 1)
  {
    for (j in 2:i)
    {
      if (mtx[1,j-1] > mtx[1,j])
      {
        t1 = mtx[1,j-1];
        t2 = mtx[2,j-1];
        t3 = mtx[3,j-1];
        mtx[1,j-1] = mtx[1,j];
        mtx[2,j-1] = mtx[2,j];
        mtx[3,j-1] = mtx[3,j];
```

```
        mtx[1,j] = t1;
        mtx[2,j] = t2;
        mtx[3,j] = t3;
      }
    }
    i = i - 1;
  }

  if (snd == TRUE)
  {
    t=mtx[1,];
    mtx[1,] = mtx[2,];
    mtx[2,] = t;
  }
  return(mtx);
}

############################################################
# \int G_m dH^1_N dH^2_N  is  EmpIntegralSpearman(X,Y)
# \int F_n dH^1_N dH^2_N  is  EmpIntegralSpearman(Y,X)
# Hence Wnm = sqrt(nm/n+m) *
#   ( EmpIntegralSpearman(X,Y) - EmpIntegralSpearman(Y,X) )
# \int H_N(u1,u2) dH^1_N(u1) dH^2_N(u2) is
# EmpIntegralSpearman(X,Y, TRUE)
############################################################
EmpIntegralSpearman<-function(X, Y, combined = FALSE)
{
    ret = 0;
```

```
n = dim(X)[2];
m = dim(Y)[2];
N = n+m;
#print("N");
#print(N);
Z = cbind(X,Y);
ZRanks = matrix(,3,N);
ZRanks[1,] = rank(Z[1,]);
ZRanks[2,] = rank(Z[2,]);
#print(ZRanks);
for (i in 1:n)
{
    if (combined == TRUE)
    {
        ZRanks[3,i] = 1;
    }
    else
    {
        ZRanks[3,i] = 0;
    }

}
for (i in 1:m)
{

    ZRanks[3,i+n] = 1;
}
```

```
    ZRanks = bubbleSortSpearman(ZRanks);

    #print(ZRanks);
    for (h in 1:N)
    {
        if (ZRanks[3,h] == 1)
        {
            ret = ret+(N-ZRanks[1,h]+1)*(N-ZRanks[2,h]+1)
        }
    }
    if (combined == TRUE)
    {
        ret = ret/(N^3);
    }
    else
    {
        ret = ret/(m*N^2);
    }

    return(ret);
}

############################################################
# Estimate the variance of Spearman statistic
############################################################
EstimateSigmaSqSpearman<-function(X,Y)
{
```

```
n = dim(X)[2];
m = dim(Y)[2];
N = n+m;
Z = cbind(X,Y);
ZRanks = matrix(,2,N);
ZRanks[1,] = rank(Z[1,]);
ZRanks[2,] = rank(Z[2,]);

# Notice that here we use bubbleSort instead of bubbleSortSpearman
# This is because ZRanks = matrix(,2,N) not ZRanks = matrix(,3,N);
ZRanks = bubbleSort(ZRanks);
#print(ZRanks);

# This is the square root of second part of the variance estimate.
sndPart = EmpIntegralSpearman(X,Y, TRUE);

# This value has been checked with the true value. Good.
sndPart = sndPart^2;
#print(sndPart);
ret = 0;
for (l in 1:N)
{
    ret = ret + (N-ZRanks[1,l]+1)^2*(N-ZRanks[2,l]+1)^2;
}
#print(ret/N^5);
ret = ret/N^5 - sndPart;

return(ret);
```

```
}
############################################
# This is the actual function to
# perform the Spearman statistic
# returns statistic and p-value
############################################
SpearmanStat<-function(X,Y)
{
   n = dim(X)[2];
   m = dim(Y)[2];
   Wnm = sqrt((n*m)/(n+m)) *  (EmpIntegralSpearman(X,Y) - EmpIntegralSpearman(Y,X)
   sig =  sqrt(EstimateSigmaSqSpearman(X,Y));
   stat = Wnm/sig;
   pval = pnorm(stat, 0, 1, lower=FALSE);
   return(c(stat,pval));

}
```

A.3 R-Code Section 4.2

A.3.1 R-code for shifted distributions

The following code did the following scenarios:

1. When $H_0 : F = G$ is true: let X, Y be iid with distribution $N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 10 & 3 \\ 3 & 2 \end{pmatrix})$. We calculate the three test statistics and compute their p-values.

2. When $H_1 : F < G$ is true: let $X1$ have distribution $N(\binom{1}{1}, \begin{pmatrix} 10 & 3 \\ 3 & 2 \end{pmatrix})$; $Y1$ have distribution $N(\binom{0}{0}, \begin{pmatrix} 10 & 3 \\ 3 & 2 \end{pmatrix})$. We calculate the three test statistics and compute their p-values.

The following code is for sample size $n = m = 200$; the sample size can be changed to any number.

```
########################################################
# Two Tests Z and Z1 in multivariate normal using M-W-W
# Kendall and Spearman statistics
# where the covariance is Sigma
# Z and Z1 are composed with X and Y, X_1 and Y_1
# This is for sample size n=m=200.
########################################################
Sigma <- matrix(c(10,3,3,2),2,2);
X = mvrnorm(n=200, rep(0,2), Sigma);
X1 = mvrnorm(n=200, rep(1,2), Sigma);
X = t(X);
X1 = t(X1);
#X;
Y = mvrnorm(n=200, rep(0,2), Sigma);
Y1 = mvrnorm(n=200, rep(0,2), Sigma);
Y = t(Y);
Y1 = t(Y1);
#Y;
Z = cbind(X,Y);
Z1 = cbind(X1,Y1);
mvstat = MVManWilTest(Z, 2, 200, 200);
mvstat1 = MVManWilTest(Z1, 2, 200, 200);
```

```
mvstat;
mvstat1;
pchisq(mvstat, 2, 0, FALSE);
pchisq(mvstat1, 2, 0, FALSE);
# when H_0 is true
KendallStat(X,Y);
SpearmanStat(X,Y);
# when H_1 is true
KendallStat(X1,Y1);
SpearmanStat(X1,Y1);
```

A.3.2 R-code for equal marginals, different copulas

The following code is for sample size $n = m = 200$; the sample size can be changed to any number.

```
##############################################################
# Compare M-W-W, Kendall and Spearman statistics
# using
#   (1)\theta_X=-1, \theta_Y=-0.5
#   (2)\theta_X=-1, \theta_Y=0.5
#   (3)\theta_X=-1, \theta_Y=1
##############################################################
X = FGMSimulation(200, -1);
Y = FGMSimulation(200, -0.5);
Z = cbind(X,Y);
mvstat = MVManWilTest(Z, 2, 200, 200);
mvstat; # M-W-W statistic
pchisq(mvstat, 2, 0, FALSE); # M-W-W p-value
```

```
KendallStat(X,Y);
SpearmanStat(X,Y);

##############################################
X = FGMSimulation(200, -1);
Y = FGMSimulation(200, 0.5);
Z = cbind(X,Y);
mvstat = MVManWilTest(Z, 2, 200, 200);
mvstat; # M-W-W statistic
pchisq(mvstat, 2, 0, FALSE); # M-W-W p-value
KendallStat(X,Y);
SpearmanStat(X,Y);

##############################################
X = FGMSimulation(200, -1);
Y = FGMSimulation(200, 1);
Z = cbind(X,Y);
mvstat = MVManWilTest(Z, 2, 200, 200);
mvstat; # M-W-W statistic
pchisq(mvstat, 2, 0, FALSE); # M-W-W p-value
KendallStat(X,Y);
SpearmanStat(X,Y);
```

A.3.3 R-code for different marginals, different copulas

```
##############################################################
# Compare M-W-W, Kendall and Spearman statistics
```

```
# using \theta_X=-1, \theta_Y=1 and
#   (1)F_1, F_2~exp(1)   G_1~exp(1), G_2~exp(1.75)
#   (2)F_1, F_2~exp(1)   G_1~exp(1), G_2~exp(1.5)
#   (3)F_1, F_2~exp(1)   G_1~exp(1), G_2~exp(1.25)
#   (4)F_1, F_2~exp(1)   G_1~exp(1), G_2~exp(1.125)
##############################################################
#### (1)F_1, F_2~exp(1)   G_1~exp(1), G_2~exp(1.75) #####
U = FGMSimulation(200, -1);
V = FGMSimulation(200, 1);
X=-log(1-U);
Y = matrix(,2,200);
Y[1,] = -log(1-V[1,]);
Y[2,] = -log(1-V[2,])/1.75;
Z=cbind(X,Y);
stat=MVManWilTest(Z, 2, 200, 200);
stat;
pchisq(stat, 2, 0, FALSE);
KendallStat(X,Y);
SpearmanStat(X,Y);

#### (2)F_1, F_2~exp(1)   G_1~exp(1), G_2~exp(1.5) #####
U = FGMSimulation(200, -1);
V = FGMSimulation(200, 1);
X=-log(1-U);
Y = matrix(,2,200);
Y[1,] = -log(1-V[1,]);
Y[2,] = -log(1-V[2,])/1.5;
Z=cbind(X,Y);
```

```
stat=MVManWilTest(Z, 2, 200, 200);
stat;
pchisq(stat, 2, 0, FALSE);
KendallStat(X,Y);
SpearmanStat(X,Y);

#### (2)F_1, F_2~exp(1)   G_1~exp(1), G_2~exp(1.25) #####
U = FGMSimulation(200, -1);
V = FGMSimulation(200, 1);
X=-log(1-U);
Y = matrix(,2,200);
Y[1,] = -log(1-V[1,]);
Y[2,] = -log(1-V[2,])/1.25;
Z=cbind(X,Y);
stat=MVManWilTest(Z, 2, 200, 200);
stat;
pchisq(stat, 2, 0, FALSE);
KendallStat(X,Y);
SpearmanStat(X,Y);
```

Bibliography

[1] Hettmansperger, T.P. (1984) *Statistical Inference Based on Ranks.* John Wiley & Sons, U.S.A.

[2] Nelsen, R.B. (1999). *An Introduction to Copulas.* First Edition, Springer-Verlag, U.S.A.

[3] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes.* First Edition, Springer-Verlag, U.S.A.

[4] Jordan, A and Ivanoff,B.G. (2011) *Multidimensional P-P Plots and Precedence Tests for Point Processes on* $\Re^d$. Preprint

[5] Billingsley, P. (1999) *Convergence of probability measures.* second edition. Wiley series in Probability and Statistics.

[6] Joe, H. (1997) *Multivariate models and dependence concepts.* Chapman & Hall.

[7] Nikitin, Y. (1995) *Asymptotic Efficiency of Nonparametric Tests.* Cambridge University Press.